

Supplemental Material for

A combined functional annotation score for non-synonymous variants

by Margarida C. Lopes, Chris Joyce, Graham R. S. Ritchie, Sally L. John, Fiona Cunningham, Jennifer Asimit, Eleftheria Zeggini.

Figure S1. Receiver operating characteristic (ROC) curve of different combinations of PolyPhen-2, SIFT, PANTHER, CAROL and GERP annotation tools. CAROL makes a larger number of correct predictions (true positives, y-axis) for a given number of errors (false positives- x-axis) compared with SIFT, PolyPhen-2, PANTHER, CAROL+GERP, CAROL+PANTHER and CAROL+PANTHER+GERP, followed by CAROL+PANTHER. The area under the curve is 0.836, 0.821, 0.782, 0.852, 0.833, 0.849 and 0.834 for PolyPhen-2, SIFT, PANTHER, CAROL, CAROL+GERP, CAROL+PANTHER and CAROL+PANTHER+GERP, respectively.

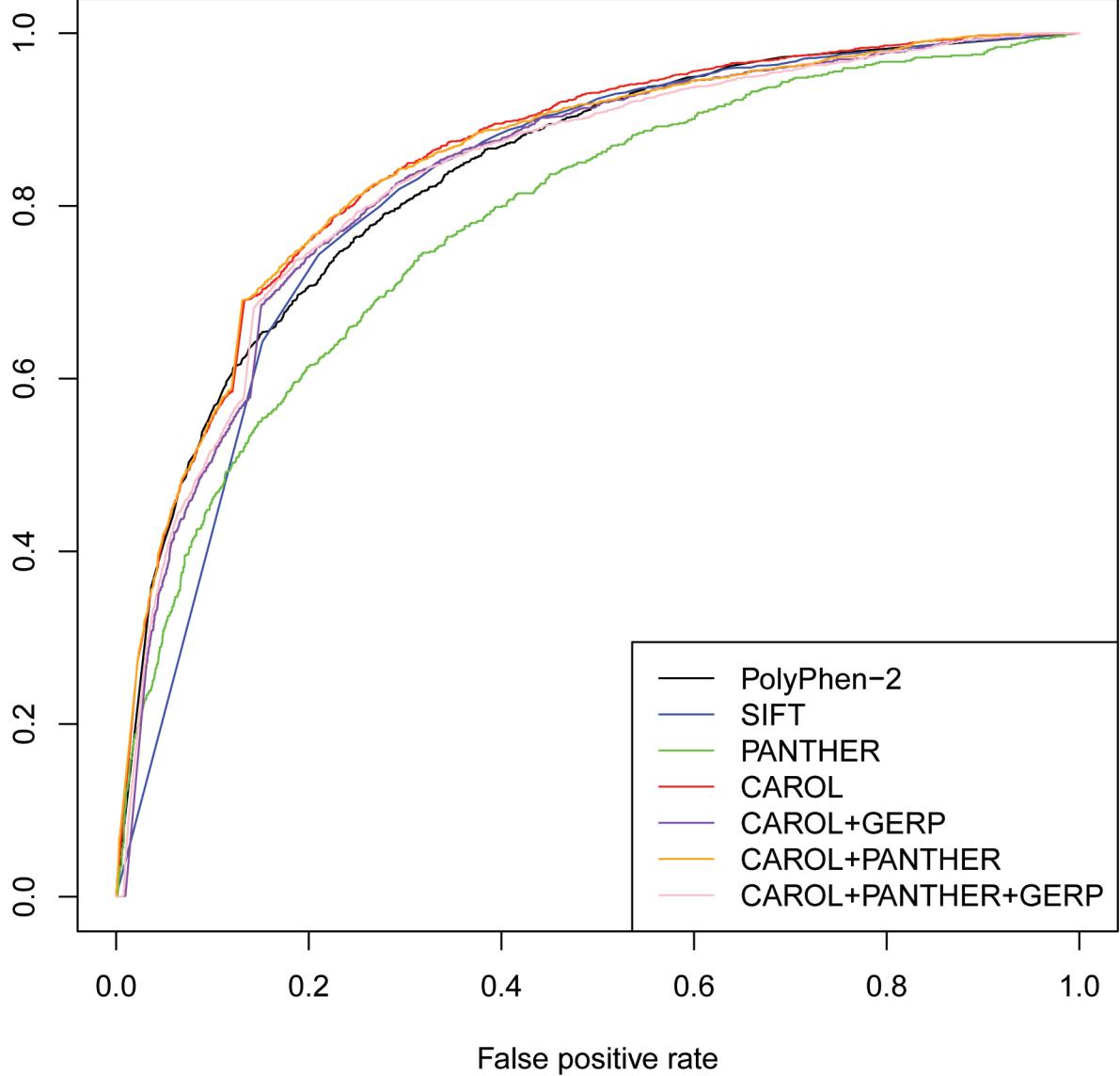
Figure S2. Probability distribution of the P_k for positive (top row graphs) and negative (bottom row graphs) controls in the test set, for each annotation tool: PolyPhen-2, SIFT and CAROL.

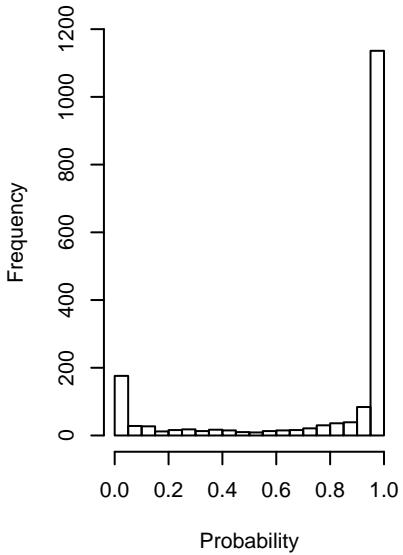
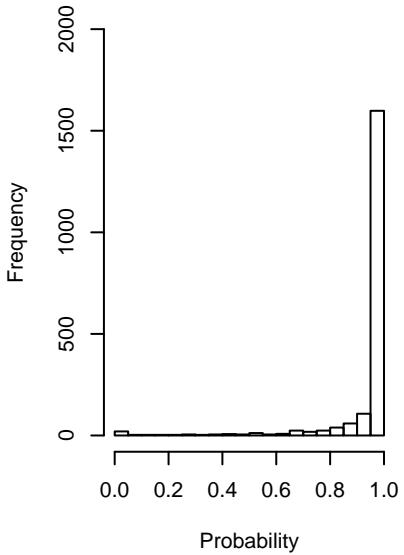
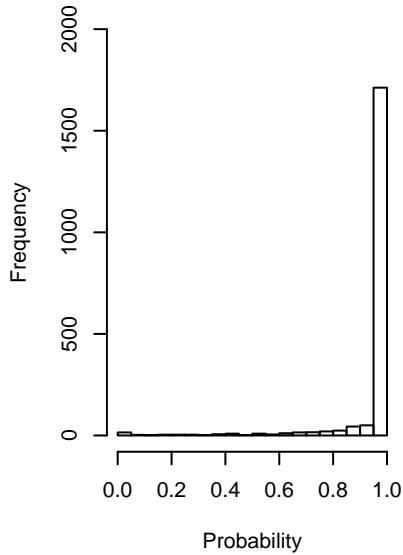
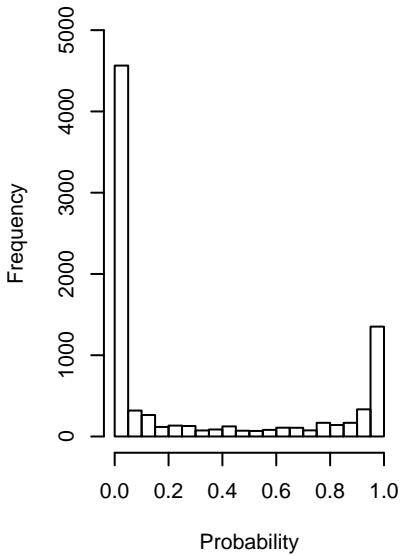
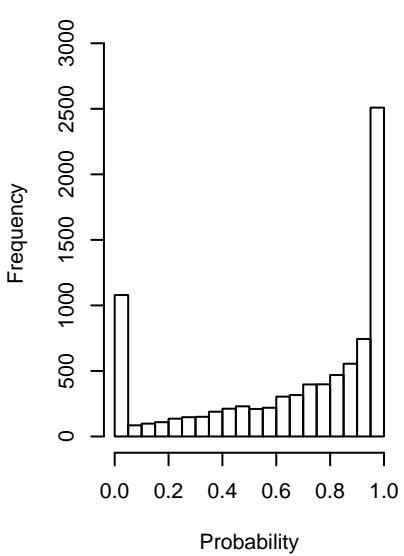
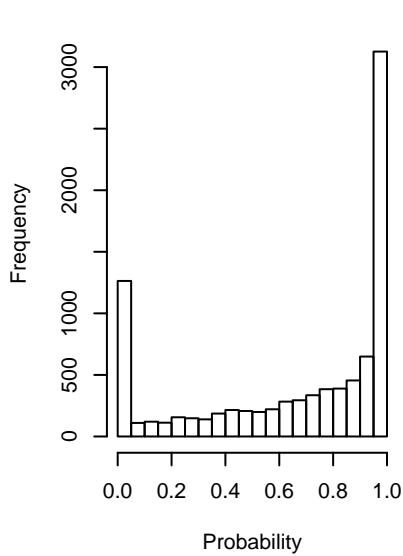
Figure S3. ROC curves generated for four annotation tools: PolyPhen-2, SIFT, CAROL and CONDEL.

Table S1. Performance statistics for the training and total datasets for the respective optimal thresholds.

	Training set			Combined set		
	PolyPhen-2	SIFT	CAROL	PolyPhen-2	SIFT	CAROL
Sensitivity	0.802	0.807	0.831	0.800	0.815	0.830
Specificity	0.707	0.732	0.740	0.708	0.728	0.731
Type II error	0.198	0.193	0.169	0.200	0.185	0.170
Type I error	0.293	0.268	0.260	0.292	0.272	0.269
NAs	0.131-0.132	0.007-0.117	0.005-0.067	0.122-0.127	0.005-0.117	0.002-0.071
Total accuracy	0.723	0.746	0.756	0.723	0.744	0.749
ROC area	0.832	0.815	0.846	0.834	0.819	0.849

True positive rate



PolyPhen-2**SIFT****CAROL****PolyPhen-2****SIFT****CAROL**

True positive rate

